

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 718 823 A2

(12)

## EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:  
26.06.1996 Patentblatt 1996/26

(51) Int. Cl. 6: G10L 9/20

(21) Anmeldenummer: 95119517.1

(22) Anmeldetag: 11.12.1995

(84) Benannte Vertragsstaaten:  
AT BE CH DE DK ES FR GB GR IT LI LU SE

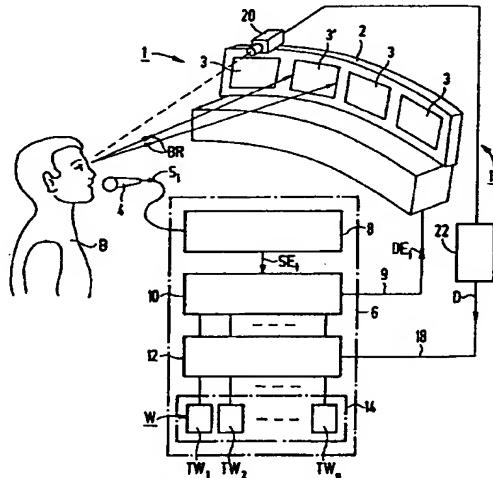
(30) Prioritat: 23.12.1994 DE 4446353

(71) Anmelder: SIEMENS AKTIENGESELLSCHAFT  
D-80333 München (DE)

(72) Erfinder:  
• Harke, Ulrike, Dipl.-Ing.  
D-83727 Schliersee (DE)  
• Mederer, Hans-Gerd, Dipl.-Phys.  
D-91052 Erlangen (DE)

### (54) Verfahren zur Umwandlung von sprachlich eingegebenen Informationen in maschinenlesbare Daten

(57) Um bei einer Umwandlung von sprachlich eingegebenen Informationen eines Benutzers (B) in ein Steuer- und Informationssystem (2) einer technischen Anlage in maschinenlesbare Daten bei hoher, ergonomischer Eingabegeschwindigkeit eine hohe Zuverlässigkeit auch bei einem komplexen zugrundeliegenden Wortschatz (W) zu gewährleisten, werden erfundungsgemäß durch einen Vergleich von aufgenommenen Sprachelementen ( $S_i$ ) mit einem anhand von Daten über den Benutzer (B) vorausgewählten Teilwortschatz ( $TW_i$ ) den Sprachelementen ( $S_i$ ) durch diese bestimmte maschinelle Datenelemente ( $DE_i$ ) zugeordnet. Durch die Auswahl eines Teilwortschatzes ( $TW_i$ ) aus einem Gesamtwortschatz (W) wird die Komplexität des zugrundeliegenden Wortschatzes (W) erst während der Spracheingabe reduziert, so daß eine hohe Erkennungsrate gewährleistet ist.



**Beschreibung**

Die Erfindung betrifft ein Verfahren zur Umwandlung von in ein Steuer- und Informationssystem einer technischen Anlage sprachlich eingegebenen Informationen eines Benutzers in maschinenlesbare Daten, bei dem durch einen Vergleich von aufgenommenen Sprachelementen mit einem vorausgewählten Teilwortschatz den Sprachelementen durch diese bestimmte maschinelle Datenelemente zugeordnet werden. Sie betrifft weiter ein System zur Spracheingabe von Informationen für eine technische Anlage.

Für den Betrieb einer technischen Anlage, insbesondere einer Kraftwerksanlage, ist eine schnelle und zuverlässige Eingabe von Informationen eines Benutzers in ein Steuer- und Informationssystem der technischen Anlage erforderlich. Für die Weiterverarbeitung in Steuer- oder Regelprozessen müssen die eingegebenen Informationen in maschinenverarbeitbare Daten, beispielsweise Codes oder Rechnerbefehle, umgewandelt werden. Insbesondere in der Leittechnik einer Kraftwerksanlage erfolgt eine derartige Informationseingabe üblicherweise über eine Tastatur oder mausunterstützt über einen Bildschirm eines Computers. Eine derartige Informationseingabe ist jedoch zeitaufwendig, insbesondere im Vergleich mit gesprochenen Informationen.

Eine schnellere Eingabe von Informationen kann mittels Spracherkennersystemen erfolgen. Dabei werden sprachlich eingegebene Informationen eines Benutzers in maschinenlesbare Daten umgesetzt, indem aufgenommene Sprachelemente mit einem Wortschatz verglichen und anhand dieses Vergleichs maschinell identifiziert werden, wobei den identifizierten Sprachelementen maschinelle Datenelemente zugeordnet werden. Übliche Spracherkennner ordnen eine erkannte Lautfolge einer Liste von Namen von Bedienelementen zu und wählen das wahrscheinlichste Listendokument aus. Bei herkömmlichen Spracherkennersystemen nimmt jedoch die Erkennungsrate der Sprachelemente und somit die Zuverlässigkeit der Spracherkennung mit zunehmender Größe des Wortschatzes stark ab. Bei komplexen Wortsätzen, wie sie insbesondere beim Betrieb einer technischen Anlage oder einer Kraftwerksanlage auftreten, sind herkömmliche Spracherkennner daher unzuverlässig und somit nicht einsetzbar. Um die Erkennungsrate eines Spracherkenners zu erhöhen, muß der zugrundeliegende Wortschatz, zum Beispiel im Hinblick auf die Anzahl der enthaltenen Wörter, eingeschränkt werden. Dies schränkt jedoch die Flexibilität eines derartigen Spracheingabesystems ein.

Ein System und ein Verfahren zur Spracheingabe von Informationen, bei dem der zugrundeliegende Wortschatz in eine Anzahl von Teilwortsätzen gegliedert ist, ist aus der DE-A-42 16 455 bekannt. Die Auswahl des Teilwortsatzes erfolgt dabei mittels eines sprachlich eingegebenen Identifizierungscodes. Eine derartige Vorauswahl ist jedoch zeitaufwendig und, insbesondere in Notfallsituationen, nur eingeschränkt zuverlässig, da

der Identifizierungscode mit hoher sprachlicher Genauigkeit einzugeben ist.

Der Erfindung liegt daher die Aufgabe zugrunde, ein Verfahren der obengenannten Art zur Umwandlung von sprachlich eingegebenen Informationen eines Benutzers in ein Steuer- und Informationssystem einer technischen Anlage in maschinenlesbare Daten anzugeben, das bei einem komplexen zugrundeliegenden Wortschatz bei hoher Eingabegeschwindigkeit eine hohe Zuverlässigkeit gewährleistet. Dies soll mit einem besonders geeigneten System zur Spracheingabe von Informationen für eine technische Anlage erreicht werden.

Bezüglich des Verfahrens wird diese Aufgabe erfindungsgemäß gelöst, indem die Vorauswahl des Teilwortsatzes anhand der räumlichen Position des Benutzers und/oder anhand der Gestik des Benutzers und/oder anhand der Blickrichtung des Benutzers und/oder anhand der Sprechrichtung des Benutzers erfolgt.

Die Erfindung geht dabei von der Überlegung aus, daß bei einer Aufteilung eines für eine technische Anlage relevanten Wortschatzes in Teilwortsätze einer der Teilwortsätze vorausgewählt werden kann. Auf der Grundlage des ausgewählten Teilwortsatzes können dann aufgenommene Sprachelemente mit Elementen des Teilwortsatzes, insbesondere mit Wörtern oder Wortbestandteilen, verglichen und somit identifiziert werden. Durch die Auswahl eines Teilwortsatzes wird die Komplexität des zugrundeliegenden Wortschatzes erst während der Spracheingabe reduziert, so daß auch bei einem komplexen zugrundeliegenden Wortschatz eine hohe Erkennungsrate gewährleistet ist. Weiterhin ausgehend von der Überlegung, daß bei einer räumlich ausgedehnten Leitwarte einer Kraftwerksanlage konzeptionell zusammenhängende Anzeigegeräte und Bedienelemente eines Teilsystems der technischen Anlage örtlich zusammengegruppiert angeordnet sind, kann bereits durch die räumliche Position des Benutzers das zu bedienende Teilsystem identifiziert und somit der diesem Teilsystem zugeordnete Teilwortschatz ausgewählt werden, wobei sich der Benutzer zur Bedienung dieses Teilsystems in dessen Nähe aufhält.

Durch die Vorauswahl des Teilwortsatzes anhand der Gestik wird dem Benutzer beispielsweise ermöglicht, durch Deuten auf ein Teilsystem der technischen Anlage, beispielsweise durch eine Armbewegung, das zu bedienende Teilsystem zu identifizieren.

Ausgehend von der Beobachtung, daß konzeptionell zusammenhängende Anzeigegeräte und Bedienelemente eines zu bedienenden Teilsystems der technischen Anlage üblicherweise in einem gemeinsamen Blickbereich des Benutzers angeordnet sind, kann zudem durch die Ermittlung der Blickrichtung des Benutzers das zu bedienende Teilsystem frühzeitig erkannt und auf dieser Grundlage eine zutreffende Auswahl des Teilwortsatzes erreicht werden.

Bezüglich des Systems zur Spracheingabe von Informationen für eine technische Anlage, bei dem ein Sprachdetektor zur Aufnahme von Sprachelementen eines Benutzers und ein Spracherkennnermodul zur

Zuordnung von maschinenlesbaren Datenelementen zu Sprachelementen vorgesehen sind, wobei das Spracherkennnermodul einen Analysebaustein zur Vorauswahl eines Teilwortschatzes aus einem im Spracherkennnermodul hinterlegten Wortschatz aufweist, wird die Aufgabe erfundungsgemäß gelöst durch ein Beobachtungsinstrument zur Beobachtung des Benutzers.

Von dem Sprachdetektor aufgenommene Sprachelemente eines Benutzers werden an das Spracherkennnermodul weitergeleitet. Durch den Analysebaustein des Spracherkennnermoduls erfolgt eine Vorauswahl eines Teilwortschatzes aus einem im Spracherkennnermodul hinterlegten Wortschatz anhand von von dem Beobachtungsinstrument ermittelten Daten über dem Benutzers. Durch Vergleich der aufgenommenen Sprachelemente mit dem somit ausgewählten Teilwortschatz werden den Sprachelementen maschinelle Datenelemente zugeordnet, wobei eine hohe Zuverlässigkeit der Spracherkennung durch Reduktion des Gesamtwortschatzes auf den Teilwortschatz gewährleistet ist.

Zur Auswahl von für die technische Anlage relevanten Schlüsselsprachelementen aus allen aufgenommenen Sprachelementen weist das Spracherkennnermodul zweckmäßigerweise einen Filterbaustein auf.

In vorteilhafter Ausgestaltung ist das Beobachtungsinstrument ein Positionsgeber zur Ermittlung der räumlichen Position des Benutzers. Anhand der mittels des Positionsgitters ermittelten Position des Benutzers kann ein zu bedienendes Teilsystem der technischen Anlage identifiziert und ein diesem Teilsystem zugeordneter Teilwortschatz ausgewählt werden.

Gemäß einer zweckmäßigen Weiterbildung ist das Beobachtungsinstrument zur Ermittlung der Blickrichtung oder der Gestik des Benutzers vorgesehen. Anhand der dadurch ermittelbaren Blickrichtung oder Gestik des Benutzers kann ebenfalls die Identifikation eines zu bedienenden Teilsystems der technischen Anlage und somit die Auswahl des diesem zugrundeliegenden Teilwortschatzes erfolgen.

Zur Ermittlung der Sprechrichtung des Benutzers weist das Spracheingabesystem einen Detektor auf.

Die mit der Erfindung erzielten Vorteile bestehen, insbesondere darin, daß durch die Auswahl eines Teilwortschatzes aus einem Gesamtwortschatz bei einem Spracheingabesystem die Zuverlässigkeit der Zuordnung von maschinellen Datenelementen zu aufgenommenen Sprachelementen auch bei einem zugrundeliegenden komplexen Gesamtwortschatz gewährleistet ist. Durch die Auswahl des Teilwortschatzes anhand von von der Spracheingabe unabhängigen Daten über den Benutzer sind ein derartiges Verfahren und ein derartiges System besonders effektiv und besonders zuverlässig.

Ein Ausführungsbeispiel der Erfindung wird anhand einer Zeichnung näher erläutert. Diese zeigt ein System 1 zur Spracheingabe von Informationen in eine Leitwarte 2 einer nicht näher dargestellten Kraftwerksanlage. An

der Leitwarte 2 sind Monitore 3 zur Darstellung von Informationen und prozeßrelevanten Daten der Kraftwerksanlage angeordnet. Die Kraftwerksanlage ist organisatorisch in Teilsysteme gegliedert, denen jeweils einer der Monitore 3 zur Darstellung der sie betreffenden Informationen zugeordnet ist.

Das Spracheingabesystem 1 umfaßt einen Sprachdetektor 4 - beispielsweise ein Mikrofon - und ein Spracherkennnermodul 6. Das Spracherkennnermodul 6 umfaßt einen Filterbaustein 8, eine über eine Datenleitung 9 mit der Leitwarte 2 verbundene Zuordnungseinheit 10, einen Analysebaustein 12 und eine Datenbank 14. In der Datenbank 14 ist ein in mehrere Teilwortschatz  $TW_i$ , mit  $i = 1, \dots, n$ , gegliederter Gesamtwortschatz  $W$  hinterlegt.

Ein Beobachtungsinstrument 16 ist über eine Leitung 18 mit dem Analysebaustein 12 verbunden. Das Beobachtungsinstrument 16 umfaßt einen Detektor 20 und eine Kontrolleinheit 22. Der Detektor 20 ist zweckmäßigerweise eine Kamera. Er kann alternativ aber auch ein Mikrofonarray, ein Headtracker oder ein Eye-tracker sein. Anstelle der Leitungen 9, 18 können auch Funkverbindungen oder andere Datenübertragungsmittel vorgesehen sein.

Mittels der Kamera 20 wird die durch den Pfeil BR angedeutete Blickrichtung eines Benutzers B des Spracheingabesystems 1 durch das Kontrollsystem 22 des Beobachtungsinstruments 16 erfaßt und in einen die Blickrichtung beschreibenden Datensatz D umgewandelt. Der Datensatz D wird über die Leitung 18 an den Analysebaustein 12 übergeben. Anhand des Datensatzes D wird im Analysebaustein 12 der vom Benutzer B angeblickte Monitor 3' ermittelt und somit das zu bedienende und diesem Monitor 3' zugeordnete Teilsystem der Kraftwerksanlage identifiziert. Ein diesem Teilsystem zugrundeliegender Teilwortschatz  $TW_i$  wird vom Analysemodul 12 aus den Teilwortschätzen  $TW_i$  des in der Datenbank 14 hinterlegten Gesamtwortschatzes  $W$  ausgewählt. Der ausgewählte Teilwortschatz  $TW_i$  wird vom Analysemodul 12 der Vergleichseinheit 10 bereitgestellt. Mittels des Sprachdetektors 4 aufgenommene Sprachelemente  $S_i$  des Benutzers B werden dem Filterbaustein 8 zugeführt. Im Filterbaustein 8 werden aus den aufgenommenen Sprachelementen  $S_i$  alle relevanten Schlüsselsprachelemente ausgewählt. Die Schlüsselsprachelemente  $SE_i$  werden der Vergleichseinheit 10 zugeführt. Anhand des ausgewählten Teilwortschatzes  $TW_i$  werden den ausgewählten Schlüsselsprachelementen  $SE_i$  maschinell verarbeitbare Datenelemente  $DE_i$  zugeordnet. Die Datenelemente  $DE_i$ , insbesondere Befehscodes, werden über die Datenleitung 9 der Leitwarte 2 zugeführt und dort zur Steuerung oder Informationsverarbeitung der Kraftwerksanlage herangezogen.

## 55 Patentansprüche

1. Verfahren zur Umwandlung von in ein Steuer- und Informationssystem (2) einer technischen Anlage sprachlich eingegebenen Informationen eines

Benutzers (B) in maschinenlesbare Daten, wobei durch einen Vergleich von aufgenommenen Sprachelementen ( $S_i$ ) mit einem vorausgewählten Teilwortschatz ( $TW_i$ ) den Sprachelementen ( $S_i$ ) durch diese bestimmte maschinelle Datenelemente ( $DE_i$ ) zugeordnet werden,  
 dadurch gekennzeichnet, daß eine Vorauswahl des Teilwortschatzes ( $TW_i$ ) aus einem hinterlegten Gesamtwortschatz (W) anhand der räumlichen Position des Benutzers (B) und/oder anhand der Gestik des Benutzers (B) und/oder anhand der Blickrichtung (BR) des Benutzers (B) und/oder anhand der Sprechrichtung des Benutzers (B) erfolgt.

5

10

15

2. System (1) zur Spracheingabe von Informationen für eine technische Anlage, insbesondere zur Durchführung des Verfahrens nach Anspruch 1, mit einem Sprachdetektor (4) zur Aufnahme von Sprachelementen ( $S_i$ ) eines Benutzers (B) und mit einem Spracherkennnermodul (6) zur Zuordnung von maschinenlesbaren Datenelementen ( $DE_i$ ) zu Sprachelementen ( $S_i$ ), wobei das Spracherkennnermodul (6) einen Analysebaustein (12) zur Vorauswahl eines Teilwortschatzes ( $TW_i$ ) aus einem im Spracherkennnermodul (6) hinterlegten Gesamtwortschatz (W) aufweist,  
**gekennzeichnet durch** ein Beobachtungsinstrument (16) zur Beobachtung des Benutzers.

30

20

25

3. Spracheingabesystem nach Anspruch 2,  
**dadurch gekennzeichnet**, daß das Beobachtungsinstrument ein Positionsgeber zur Ermittlung der räumlichen Position des Benutzers (B) ist.

35

4. Spracheingabesystem nach einem der Ansprüche 2 oder 3,  
**dadurch gekennzeichnet**, daß das Beobachtungsinstrument (16) zur Ermittlung der Blickrichtung (BR) oder der Gestik des Benutzers (B) vorgesehen ist.

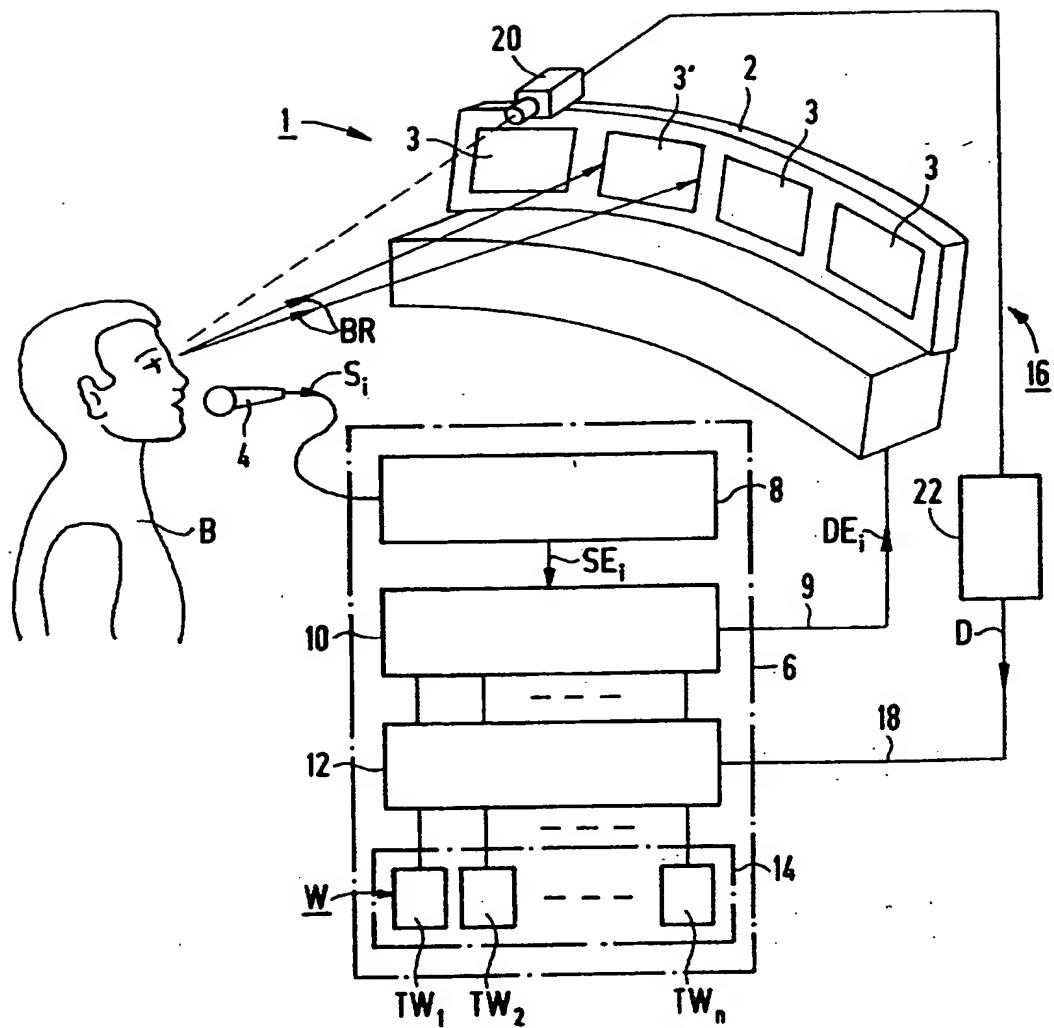
40

5. Spracheingabesystem nach einem der Ansprüche 2 bis 4,  
**gekennzeichnet durch** einen Detektor zur Ermittlung der Sprechrichtung des Benutzers (B).

45

50

55





(19) Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 718 823 A3

(12)

## EUROPÄISCHE PATENTANMELDUNG

(88) Veröffentlichungstag A3:  
14.01.1998 Patentblatt 1998/03

(51) Int. Cl. 6: G10L 9/20

(43) Veröffentlichungstag A2:  
26.06.1996 Patentblatt 1996/26

(21) Anmeldenummer: 95119517.1

(22) Anmeldetag: 11.12.1995

(84) Benannte Vertragsstaaten:  
AT BE CH DE DK ES FR GB GR IT LI LU SE

(72) Erfinder:

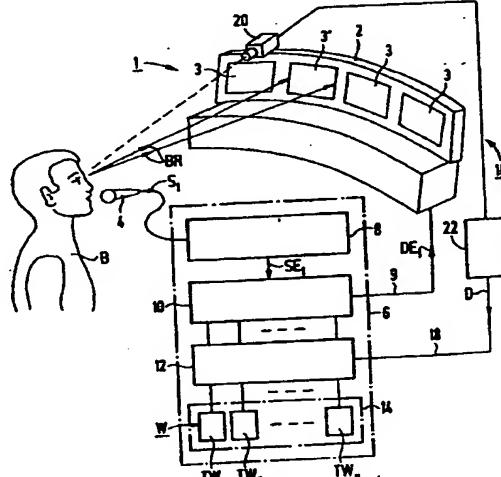
- Harke, Ulrike, Dipl.-Ing.  
D-83727 Schliersee (DE)
- Mederer, Hans-Gerd, Dipl.-Phys.  
D-91052 Erlangen (DE)

(30) Priorität: 23.12.1994 DE 4446353

(71) Anmelder:  
SIEMENS AKTIENGESELLSCHAFT  
80333 München (DE)

(54) Verfahren zur Umwandlung von sprachlich eingegebenen Informationen in  
maschinenlesbare Daten

(57) Um bei einer Umwandlung von sprachlich eingegebenen Informationen eines Benutzers (B) in ein Steuer- und Informationssystem (2) einer technischen Anlage in maschinenlesbare Daten bei hoher, ergonomischer Eingabegeschwindigkeit eine hohe Zuverlässigkeit auch bei einem komplexen zugrundeliegenden Wortschatz (W) zu gewährleisten, werden erfundsgemäß durch einen Vergleich von aufgenommenen Sprachelementen (S<sub>i</sub>) mit einem anhand von Daten über den Benutzer (B) vorausgewählten Teilwortschatz (TW<sub>i</sub>) den Sprachelementen (S<sub>i</sub>) durch diese bestimmte maschinelle Datenelemente (DE<sub>i</sub>) zugeordnet. Durch die Auswahl eines Teilwortschatzes (TW<sub>i</sub>) aus einem Gesamtwortschatz (W) wird die Komplexität des zugrundeliegenden Wortschatzes (W) erst während der Spracheingabe reduziert, so daß eine hohe Erkennungsrate gewährleistet ist.



EP 0 718 823 A3



Europäisches  
Patentamt

## EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung  
EP 95 11 9517

EINSCHLÄGIGE DOKUMENTE							
Kategorie	Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile	Betreff Anspruch	KLASSIFIKATION DER ANMELDUNG (Int.Cl.6)				
P, X	EP 0 653 701 A (IBM) * Zusammenfassung *	1	G10L9/20				
A	EP 0 472 356 A (FUJITSU TEN LTD) * Zusammenfassung *	1					
A	EP 0 618 565 A (IBM) * Zusammenfassung; Ansprüche 1-4 *	1					
<table border="1"> <tr> <td colspan="2">RECHERCHIERTE SACHGEBiete (Int.Cl.6)</td> </tr> <tr> <td colspan="2">G10L</td> </tr> </table>				RECHERCHIERTE SACHGEBiete (Int.Cl.6)		G10L	
RECHERCHIERTE SACHGEBiete (Int.Cl.6)							
G10L							
<p>Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt</p>							
Recherchenort	Abschlußdatum der Recherche	Prüfer					
DEN HAAG	20.November 1997	Van Doremalen, J					
KATEGORIE DER GENANNTEN DOKUMENTE		<p>T : der Erfindung zugrunde liegende Theorien oder Grundsätze            E : älteres Patentdokument, das jedoch erst am oder            nach dem Anmeldedatum veröffentlicht worden ist            D : in der Anmeldung angeführtes Dokument            L : aus anderen Gründen angeführtes Dokument            8 : Mitglied der gleichen Patentfamilie, übereinstimmendes            Dokument</p>					
X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur							

THIS PAGE BLANK (USPTO)